The Orthographic Expectation Effect in Oral Vocabulary Learning of Chinese L2 Learners

Rongyue Zhang[a], Tianlin Wang[b], Hui Sun[a], Miao Yu[a]

[a]Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Faculty of Psychology, Tianjin Normal University, 393 Binshui West Road, Xiqing District, Tianjin 300387, China

[b]Educational & Counseling Psychology, University at Albany, State University of New York, New York, United States of America

Corresponding author:

Miao Yu, Academy of Psychology and Behavior, Faculty of Psychology**,** Tianjin Normal University, 393 Binshui West Road, Xiqing District, Tianjin 300387**,** China

E-mail: ym@tjnu.edu.cn

Abstract: Previous research has indicated that English native speakers can develop orthographic expectancy for new words solely through auditory learning, it remains unclear whether orthographic expectancy exists in the process of learning a second language employing an ideographic writing system. This study aims to investigate whether Chinese second language (L2) learners can form orthographic expectancy for newly acquired Chinese characters across two experiments. The findings show that through auditory learning, Chinese L2 learners with higher writing proficiency can form orthographic expectancy for phonogram characters by utilizing

phonology-orthography mapping knowledge. These outcomes present the latest evidence regarding whether Chinese L2 learners can generate orthographic expectancy during Chinese character acquisition, underscoring the universality of the *orthographic skeleton hypothesis*. The pedagogical implications for vocabulary teaching in the context of learning Chinese as a second language are also discussed.

Keywords Oral Vocabulary Learning, Chinese Second Language Learners, Orthographic Expectancy, Chinese Character Types, Writing Proficiency

Introduction

Spoken language constitutes an indispensable mode of expression and communication in people's daily lives. When individuals hear or learn a spoken word, they typically acquire information regarding its pronunciation and meaning (Levelt et al., 1999). Empirical studies focusing on vocabulary learning have consistently found that spoken vocabulary facilitates reading (Duff & Hulme, 2012; McKague et al., 2001; Nation & Cocksey, 2009; Zhou et al., 2015). Given the significance of spoken vocabulary in reading, investigating how spoken language may underlie the development of reading could provide insight into language acquisition, thereby creating optimal conditions for learners in learning how to read. However, the precise mechanisms through which spoken vocabulary influences reading remain less understood. Some researchers have proposed that learners can map speech onto word forms, leading to the generation of orthographic expectations (i.e., preliminary orthographic representations or orthographic skeletons) upon encountering new words auditorily, thereby aiding reading (Jevtović et al., 2022; Stuart & Coltheart, 1988; Wegener et al., 2023). Yet, this notion has only been validated in alphabetic languages, leaving the existence of orthographic expectancy in non-alphabetic languages uncertain. Therefore, this study aims to explore whether L2 learners, during the process of learning Chinese, an ideographic writing system, can also generate orthographic expectancy solely through auditory learning, akin to alphabetic languages.

Background Literature

Recently, Wegener et al. (2018) conducted a child experiment directly verifying orthographic representation generated solely through auditory learning. They employed a novel word learning paradigm, teaching fourth-grade English native-speaking children the pronunciation and meaning

of a set of spoken words, followed by reading contextual sentences containing either trained or untrained words. Results revealed significantly shorter reading time for newly trained words through spoken language than untrained words. Additionally, for new words with predictable spellings (e.g., /neʃ/ spelled as nesh), reading time were significantly shorter compared to unpredictable spellings (e.g., /kɔb/ spelled as koyb). Crucially, words trained through spoken language exhibited a larger difference in gaze duration between predictable and unpredictable spellings compared to words untrained through spoken language. Based on these findings, researchers proposed the *orthographic skeleton hypothesis* to explicate the mechanism by which spoken vocabulary influences reading. This hypothesis posits that learners, during spoken vocabulary acquisition, can utilize their knowledge of sound-to-form mapping to generate orthographic expectancy or construct an orthographic skeleton of the word before encountering its written form. If the orthographic expectancy matches the actual orthographic form upon first sight, online processing becomes easier, resulting in shorter reading time. Conversely, mismatches between orthographic expectancy and actual written forms incur processing costs, leading to longer reading time. Subsequently, similar results were observed in adult English learners (Beyersmann et al., 2023) as well as in an isolated word recognition task (Wegener et al., 2023), further supporting the *orthographic skeleton hypothesis*.

Currently, prior research has predominantly focused on native speakers (Jevtović et al., 2022; 2023; Wegener et al., 2018). Existing studies indicate that Chinese native speakers can automatically activate relevant orthographic information during spoken language processing (Qu & Damian, 2017; Zou et al., 2012). However, the developmental process of language acquisition differs between second language learners and native speakers. On the one hand, native speakers typically first develop spoken language skills and subsequently acquire written language,

whereas L2 learners simultaneously acquire spoken and written language (Veivo et al., 2018). This parallel learning approach might exert unique influences on novel word learning (Marian et al., 2021), potentially fostering a closer association between spoken vocabulary and written forms among L2 learners. Consequently, it can be predicted that L2 learners exhibit more significant orthographic activation during auditory processing compared to native speakers. However, on the other hand, vocabulary knowledge among L2 learners might be subject to influences from learning environments, frequency of use, among other factors, rendering L2 lexical items unstable and incomplete across various representational forms (semantic, syntactic, phonological, and orthographic) (Jiang, 2000). Based on this assumption, it can also be anticipated that the orthographic activation during auditory processing among L2 learners might not be as pronounced as it is among native speakers (Qu et al., 2018). Overall, whether L2 learners can generate orthographic expectancy during spoken language learning through sound-to-form mapping skills deserve further investigation.

Moreover, the *orthographic skeleton hypothesis* has been substantiated in other alphabetic languages. Jevtović et al. (2022) and Jevtović et al. (2023) independently validated this mechanism through behavioral experiments in Spanish and French native speakers, respectively, showing evidence that learners can generate orthographic expectancy for related spoken words through auditory learning. However, while each letter or symbol in alphabetic languages represents a component of pronunciation that unfold linearly, creating a close relationship between orthography and phonology (see a review of the spelling-to-sound rule, or the grapheme-phoneme correspondence (GPC) rules in Chinese in Hsieh et al., 2021), Chinese characters' components or combinations do not adhere to such rules (Perfetti et al., 2007). Specifically, Chinese characters comprise ideographic components used to record morphemes

and syllables in a planar block-shaped symbol system, aligning more with meaning than sound (Zhang et al., 2014), thus exhibiting a more arbitrary correspondence between orthography and phonology compared to most alphabetic scripts (Marian et al., 2021). However, the presence of phonetic radicals in phonogram character (semantic–phonetic compound characters) represents phonetic information, providing similar or approximate pronunciations, significantly reducing the arbitrary nature of form-sound associations in Chinese (Zhang et al., 2011).

For example, in the case of the phonogram "清(qing1)", the left component "氵" is the semantic radical, providing a category for the meaning of the character (Tong et al., 2016), indicating it is related to water. The component on the right, "青(qing1)", is the phonetic radical, offering phonetic clues about the character, allowing readers to infer the pronunciation of "清" through its phonetic radical "青". Similarly, the character "鲭(qing1)", meaning "fish", shares the same phonetic radical "青" as "清", indicating an identical or similar pronunciation. This allows learners to predict that a character containing the "青" radical will likely have a pronunciation similar to "qing". Consequently, although Chinese characters possess semantic properties, the existence of phonogram characters aligns this type of characters somewhat with alphabetic languages.

At the same time, it is important to note that many phonograms have phonetic radicals that can stand alone as characters, with their own pronunciation and meaning. Take the character "青 (qing1)" for example again, it is a single character whose original meaning is the color blue or the color of plants. This characteristic means that the phonetic radicals of phonograms not only have a phonetic function but also an orthographic function (Zhang et al., 2014). Therefore, when learners hear a certain sound, they can accurately predict the orthography of the

corresponding phonetic radical based on the relatively fixed relationship between the syllable and its written form. This provides a basis for exploring whether phonograms might exhibit an orthographic expectation effect similar to that of alphabetic writing systems. In contrast, non-phonogram Chinese characters, such as "心(xin1)", which means "heart", diverge significantly from alphabetic writing systems by lacking phonetic information. It remains to be explored whether learners can also form orthographic expectations for this type of Chinese characters, especially when compared to phonogram characters like "清".

However, it is important to note that some studies on Chinese characters aim to obtain similar effects observed in research on alphabetic languages. Although there are certain similarities between Chinese and alphabetic languages, they are two different writing systems - even if the results sometimes appear similar, there may be significant and mechanistic differences that underlie these similarities. Specifically, although the phonetic radical of phonogram characters is functionally similar to the phoneme-grapheme correspondence in alphabetic languages (Hsiao & Liu, 2010), the phoneme-grapheme correspondence unit in Chinese characters is the syllable and grapheme (McBride, 2016). As an essentially visual language, the perception of the shape of individual Chinese characters is more prominent during memory encoding (Yan et al., 2015). As we examine orthographic expectations, it is important to consider that the primary challenge in learning Chinese, especially for L2 learners, is the difficulty in writing characters. Studies on L2 learners have found significant differences among learners at different writing proficiency levels. For example, in an analysis of Chinese L2 learners' writing errors during a dictation task, Liang (2019) found that beginners had relatively low writing accuracy but showed a significant increase in accuracy as their Chinese proficiency improved. Accordingly, it can be reasonably inferred that lower proficiency L2 learners possess

less accurate and stable vocabulary representations, but as their L2 proficiency increases, vocabulary representations become more accurate (Darcy et al., 2013), resulting in more precise orthographic expectations. Chinese as a second language (CSL) learners' writing proficiency may therefore affect their orthographic expectations.

The Present Study

This study will employ a novel word learning paradigm and a lexical recognition task across two experiments to comprehensively investigate whether L2 learners of Chinese whose native languages are alphabetic languages can generate orthographic expectancy during spoken vocabulary learning. Experiment 1 will manipulate *Chinese character type* and *training condition* to explore whether L2 learners can generate orthographic expectancy for different types of Chinese characters. Due to the difference in phonetic information provided by phonograms and non-phonograms, Experiment 1 proposes the following hypothesis:

1. CSL learners will be able to form orthographic expectations for phonogram characters, but will not be able to form such expectations for non-phonogram characters.

Experiment 2 manipulates *phonetic radical predictability*, *training condition,* and *character writing proficiency* to examine whether CSL learners' orthographic expectancy for phonogram characters is impacted by the predictability of phonetic radicals as well as the learners' proficiency level for writing characters. Given that phonetic radicals in phonogram characters can also be predicted through the varying degrees of syllable-written forms correspondence among different syllables and their corresponding orthography, and CSL learners with higher writing proficiency typically excel in writing Chinese characters or possess more precision and

familiarity with regard to the written representations of Chinese characters (Chen et al., 2019; Liang, 2019). Experiment 2 proposes the following hypotheses:

2. CSL learners will be able to form orthographic expectations for phonogram characters with predictable phonetic radicals;
3. CSL learners with higher writing proficiency are able to form more precise orthographic expectations in writing tasks compared to those with lower writing proficiency.

Experiment 1: The Influence of Chinese Character Types on Orthographic Expectancy

Method

Participants

Using G*Power 3.1.9.2 software to compute the required sample size for the study (Faul et al., 2007), with an effect size set at 0.25 and α at 0.05, the calculation indicated a need for 24 participants to achieve a statistical power of 0.80. Thirty-six non-Chinese students from a university in northeastern China (mean age = 23.64 ± 3.46 years; 7 males, 29 females) were recruited for the experiment. The participants' first languages were primarily Vietnamese, Thai, and English, all of which are alphabetic languages. All participants passed the standardized *Hanyu Shuiping Kaoshi* Level 4 (Chinese proficiency test Level 4, referred to as HSK level 4), had normal or corrected normal vision, provided informed consent before the experiment, and received compensation upon completion.

Materials

To control for participants' existing learning experiences, pseudowords were used for both non-phonogram and phonogram characters. The non-phonogram characters were adapted from Tseng et al. (2023). Twenty novel non-phonogram characters were constructed in a way that conformed to orthographic conventions based on the stroke patterns of Chinese characters. Phonogram characters were selected based on the HSK levels. To ensure familiarity with the phonetic radicals, they were all from HSK Levels 1-2. Simultaneously, semantic radicals that cannot independently form characters (e.g., "彳") were chosen to avoid their semantic effects on the experimental results. Forty Chinese characters were randomly divided into two groups, A and B, each containing 10 non-phonogram and 10 phonogram characters. The mean stroke counts for the two groups were 6.80 (SD = 2.63) and 6.55 (SD = 2.33), respectively, with no significant difference between them $t(38) = 0.32$, $p = 0.752$. All pronunciation information from Chinese characters is selected from commonly used Chinese syllables; their audio stimuli were obtained from audio files provided by Baidu Chinese, voiced by a female speaker.

Forty visual images, all within semantic categories familiar to the participants (such as plants or animals), were chosen for participants to learn. To ensure semantic transparency of the images, a pretest was conducted with 20 native Chinese speakers using a semantic category judgment task (rated on a 5-point scale, 1 = "completely unrelated," 5 = "completely related"), and only images with ratings above 4 were selected. Forty images were randomly paired with a Chinese character and its pronunciation to create learning materials with visual, auditory, and semantic associations, which remained consistent for all participants.

Procedure

Each participant was individually tested in a soundproofed lab environment. The presentation of all stimuli and data collection were carried out using E-prime 3.0. The entire experiment consisted of two phases: spoken language learning and testing, with spoken language learning further divided into novel Chinese character learning and picture selection task.

Prior to learning, participants were randomly assigned to two groups, one for learning Material A (18 participants) and the other for learning Material B (18 participants). For participants assigned to Material A, only the characters in A were presented during the learning phase and constituted the *trained* condition; Material B characters were withheld as the *untrained* condition for these participants. The assignment of *trained* versus *untrained* conditions was reversed for those assigned to Material B, such that only characters from B were shown during learning. Participants were informed that they would subsequently learn some "new Chinese characters" and were asked to memorize the pronunciation and the representative image for each novel character as much as possible. During the learning phase, 20 images representing novel characters sequentially appeared at the center of the screen in a pseudorandom order, accompanied by the pronunciation of the character (e.g., During the learning process, participants could only see the picture of a "shell" and hear the pronunciation "/nian2/"), yet the written form was not displayed. Each image was presented four times, with the character's pronunciation played each time. Participants progressed to the next image at their own pace by pressing a key. An example of the materials is illustrated in Figure 1.

Figure 1 Example of learning materials.

Upon completion of the learning phase, participants underwent a picture selection task to reinforce their previously learned content. In this task, while a novel character's pronunciation was randomly played, two different images were simultaneously presented on the left and right sides of the computer screen. Participants were asked to press the "F" key (representing the left image) or the "J" key (representing the right image) to choose the image corresponding to the pronunciation. Each image appeared four times for a total of 80 trials, as illustrated in Figure 2. Immediate feedback, indicating whether the choice was correct or incorrect, was provided after each selection to reinforce learning. Upon completion of all selection tasks, an overall accuracy rate was presented. Participants whose overall accuracy rate reached 85% or higher proceeded to the testing phase, while others would undergo additional learning.

Figure 2 Example of picture selection task.

Before entering the testing phase, all 20 novel characters' pronunciations and corresponding images were replayed to help participants recall and reinforce the learned content. Participants used the "Enter" key to proceed to the next item after reviewing. Following the review, they engaged in a lexical recognition task. In this task, a fixation cross ("+") appeared at the center of the computer screen for 500 milliseconds. Subsequently, written forms of several novel characters were displayed one at a time following a pseudo-randomized order. Some characters were learned during the spoken language learning phase, while others were not. Participants were instructed to quickly judge whether the presented Chinese characters were ones

they had learned during the spoken language learning phase by pressing the "F" key for "trained" or the "J" key for "untrained." Assignment of response keys was counterbalanced among participants. Each judgment was followed by a blank screen for 1000 milliseconds. The specific procedure is illustrated in Figure 3. Before the formal start, participants underwent three practice trials differing from the actual materials to familiarize themselves with the task flow. The entire experiment lasted approximately 35 minutes. Considering the differences between Chinese characters (which do not always provide reliable phonetic clues) and alphabetic languages, only response times (recorded in milliseconds) were extracted as the dependent variable from this task.

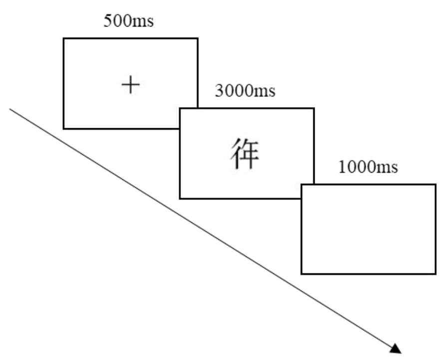

Figure 3 Flow of the lexical recognition task.

Statistical Analyses

For reaction time analysis, data beyond three standard deviations were removed (accounting for 3.13% of the raw data). To ensure linearity, reaction time data underwent log transformation. All analyses were conducted using the lme4 package (Bates et al., 2018) in the R environment (R Development Core Team, 2021) employing linear mixed models (LMMs). Training Condition, Chinese Character Type, and their interactions were included as fixed factors in the model, while participants and items were considered as random crossed factors (Baayen et al., 2008). All models began with a maximized random structure, including both intercepts and slopes (Barr et al., 2013). If a complex model failed to converge, we sequentially removed the correlations and

slopes for items, and if required, continued with those for participants until convergence was obtained. p-values for fixed effects were calculated using the lmerTest package (Kuznetsova et al., 2017).

Table 1 presents the average reaction times for different types of Chinese characters under different training conditions, and Table 2 provides the results of the linear mixed models for reaction time.

Table 1 Average reaction times (ms) for different types of Chinese characters under the training conditions

Training Condition	Chinese Character Type	
	Phonogram	Non-Phonogram
Trained	1219 (580)	1262 (562)
Untrained	1274 (569)	1240 (579)

(*Note*: Reaction times are in milliseconds, values in parentheses indicate standard deviations.)

Table 2 Results of linear mixed models for reaction times of different types of Chinese characters under the training conditions

Effect	Reaction Time			
	b	*SE*	*t*	95%CI
Intercept	7.03	0.05	138.31***	[6.93, 7.13]
Phonogram vs. Non-Phonogram	0.01	0.02	0.21	[−0.04, 0.05]
Untrained vs. Trained	0.01	0.02	0.50	[−0.03, 0.05]
Character Type × Training	0.07	0.04	1.95	[0.00, 0.15]

(*Note*: *p<0.05, **p<0.01, ***p<0.001. CI=Confidence Interval. The same applies to the following tables.)

Results

Participants demonstrated a high overall accuracy rate in the picture selection task: 98.4% (SD = 0.02), indicating strong learning of the training materials. There was no significant difference in accuracy between the two groups of learning Materials A (M = 98.3%, SD = 0.02) and B (M = 98.5%, SD = 0.02), $t(34) = -0.24$, p = 0.811.

Statistical analysis (as shown in Figure 4) in lexical recognition task indicated that the main effect of Chinese character type was not significant, $t = 0.21$, $p = 0.838$; the main effect of training condition was not significant, $t = 0.50$, $p = 0.619$; however, there was a marginally significant interaction between character type and training condition, $t = 1.95$, $p = 0.052$. Follow-up analyses revealed that only for phonograms, the reaction times of trained characters were marginally shorter than those for untrained characters, $t = -1.72$, $p = 0.085$. However, there was no significant difference in reaction time for trained versus untrained non-phonogram characters, $t = 1.03$, $p = 0.303$. These results suggest that no evidence was found under the current conditions that learners can form orthographic expectations for non-phonogram characters, but there was a trend towards forming orthographic expectations for phonogram characters following learning.

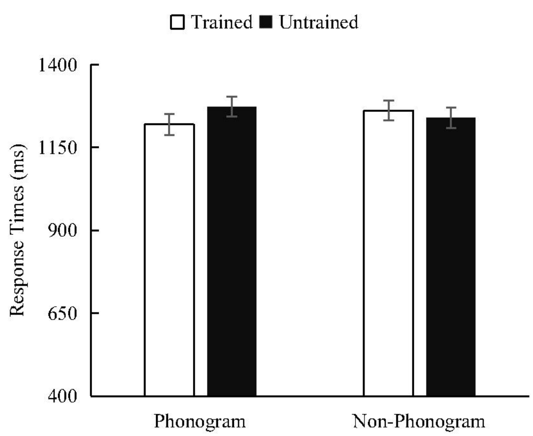

Figure 4 Average reaction times for different types of Chinese characters under the training conditions.

Discussion

The purpose of Experiment 1 was to investigate whether Chinese L2 learners could generate orthographic expectancy for different types of Chinese characters. The results indicated that there was no significant difference in reaction time between the trained and untrained conditions

for non-phonogram characters. However, for phonogram characters, the reaction time for the trained condition was marginally shorter than that for the untrained condition.

The results of Experiment 1 suggest that whether Chinese L2 learners can generate orthographic expectancy through spoken language learning might be related to the characteristics of different types of Chinese characters. Since non-phonogram characters cannot provide useful phonetic information from their written forms, learners may not predict their written forms through phonology-orthography correspondence knowledge. Hence, there was no significant difference between the trained and untrained conditions. However, phonogram characters' phonetic radicals can provide phonetic clues about the character, enabling learners to anticipate them through phonology-orthography relationships, resulting in faster reaction times for recognizing trained phonogram characters. Previous studies have shown that phonetic radicals' cues are effectively extracted and utilized by both native speakers (He et al., 2005; Yum et al., 2014) and L2 learners (Zhang & Roberts, 2019; Zhang et al., 2021). Research on the development of phonetic radical awareness in L2 students also suggests that with an increase in vocabulary knowledge, learners gradually develop phonetic radical awareness (Zhang & Roberts, 2019). Participants in this study had intermediate Chinese proficiency (HSK level 4+), indicating that they have developed phonetic radical awareness. However, experiment 1 found that the reaction times between trained and untrained phonogram characters was trending significant (p = 0.085). There may be two possible reasons for this situation.

Firstly, possibly due to the unmanipulated predictability of the phonetic radicals. Previous research indicated that reaction time are significantly shorter for predictable spellings than unpredictable spellings (Wegener et al., 2023), suggesting that predictability affects recognition time. Therefore, it can be inferred that the reaction time required to recognize

phonogram characters with the orthography of predictable phonetic radical is significantly shorter than unpredictable ones. Secondly, it may be related to the level of Chinese writing proficiency of the L2 speaker. As mentioned in the introduction, unlike alphabetic languages, one Chinese character maps onto a single syllable (McBride, 2016). This difference in the unit of mapping of phonology-orthography correspondence might imply that learning Chinese involves not merely mastering symbols but accumulating extensive vocabulary and grasping the construction of Chinese characters to comprehend phonology-orthography relationships within characters. The high visual complexity of Chinese orthography might render their learning more challenging (Ye & McBride, 2022). Consequently, learners lacking Chinese writing abilities or familiarity with the graphical representations of Chinese characters may struggle to accurately comprehend their construction and orthographic rules, hindering their ability to predict the written forms of newly trained characters based on phonology-orthography relationships. Therefore, experiment 2 further manipulates the predictability of phonetic radicals and the writing proficiency of CSL learners to explore their influence on the formation of orthographic expectancies for phonogram characters.

Experiment 2: The influence of phonetic radical predictability and L2 writing proficiency on orthographic expectancy for phonogram characters

Method

Participants

The estimated sample size for experiment 2, similar to that of experiment 1, indicated a requirement of 24 participants to achieve a statistical power of 0.80. Thirty-six non-Chinese students from a university in northeastern China were recruited (mean age = 23.64 ± 3.46, 7

males, 29 females) were recruited for the experiment. The participants' first languages were primarily Vietnamese, Thai, and English, all of which are alphabetic languages. All participants used Chinese as a second language and had passed the HSK level 4 examination at the time of testing. In the writing task evaluation, scores ranged from 0 to 20 points (M = 12.28, SD = 6.39). Based on the mean score, participants were divided into two groups: the low writing proficiency group, consisting of 16 individuals (M = 6.50, SD = 4.76), and the high writing proficiency group, consisting of 20 individuals (M = 16.90, SD = 2.59). All participants had normal vision or corrected-to-normal vision.

Materials

To ensure the orthographic predictability of selected phonetic radicals, a pretest was conducted with 20 native Chinese speakers who transcribed the radicals based on what they heard. A consistency rating of the phonetic radicals they produced was then calculated. Ultimately, 20 phonetic radicals with a consistency of 0.9 or above were selected as predictable phonetic radicals, and 20 phonetic radicals with 1 to 3 homophones were selected as unpredictable phonetic radicals (for example, hearing /fen1/ and writing "分，芬，纷" indicates a predictable phonetic radical; hearing /yue4/ and writing "月， 乐， 岳" indicates an unpredictable phonetic radical). The pronunciation of phonogram characters containing predictable phonetic radicals matched the that of the phonetic radical, while the pronunciation of those containing unpredictable phonetic radicals did not. Both types of radicals were selected from HSK level 1-2 vocabulary to ensure participants' familiarity.

Next, 40 phonogram characters were randomly divided into two groups, A and B, each containing 10 phonogram characters with predictable phonetic radicals and 10 with

unpredictable phonetic radicals. The average stroke count for the two groups was 8.05 (*SD* = 1.50) and 7.80 (*SD* = 1.96), respectively, with no significant difference between them, $t(38) = 0.45$, $p = 0.654$. Audio stimuli for all characters were obtained from Baidu Chinese. Similar to what was done in Experiment 1, 20 visual images were selected based on ratings from 20 native Chinese speakers (rated above 4.5) and paired randomly with the phonogram characters to create learning materials.

Procedure

The auditory learning phase mirrored that of Experiment 1. The testing phase included a writing task and a lexical recognition task. In the writing task, the computer screen again presented the pronunciation of new characters along with corresponding images, prompting participants to write the newly trained characters. On average, participants took approximately 10 minutes to complete the writing task before proceeding to the lexical recognition task.

Data Analysis

The data screening criteria and analyses were akin to Experiment 1. Following the removal of invalid data (accounting for 3.61% of the raw data), data analysis was conducted. Participants' average reaction times for different types of items under different training conditions are presented in Table 3, with statistical results outlined in Table 4.

Table 3 Mean reaction times (ms) of phonetic radical predictability under different training conditions

Training Condition	Phonetic Radical Predictability	
	Predictable	Unpredictable
Trained	1098 (517)	1320 (593)
Untrained	1317 (585)	1265 (570)

Table 4 Results of linear mixed model for reaction times of phonetic radical predictability under different training conditions

Effect	Reaction Time			
	b	*SE*	*t*	95%CI
Intercept	7.04	0.05	132.09^{***}	[6.93, 7.14]
Unpredictable vs. Predictable	0.07	0.03	2.17^{*}	[0.01, 0.13]
Untrained vs. Trained	0.07	0.02	3.00^{**}	[0.02, 0.11]
Predictability ×Training	−0.22	0.03	-6.70^{***}	[−0.29, −0.16]

Results

Picture selection task

Participants exhibited high overall accuracy in the picture selection task, with an accuracy rate of 98.3% (SD = 0.02), indicating strong learning of the training materials. The accuracy for learning both sets of materials (Set A: M = 98.3%, SD = 0.02; Set B: M = 98.3%, SD = 0.01) did not exhibit a significant difference, $t(34) = 0.00$, $p = 1.000$.

The Impact of Phonetic Radical Predictability on Orthographic Expectancy for Phonogram Characters

The statistical analysis results (depicted in Figure 5) in lexical recognition task revealed significant main effects of phonetic radical predictability, $t = 2.17$, $p = 0.035$, and training condition, $t = 3.00$, $p = 0.005$. Specifically, recognizing characters with predictable phonetic radical yielded significantly shorter reaction time compared to characters with unpredictable phonetic radical. Additionally, reaction times for characters trained during the experiment were notably shorter than those untrained. Importantly, a significant interaction emerged between phonetic radical predictability and training, $t = -6.70$, $p < 0.001$. Follow-up analyses indicated that, under the *trained* condition, reaction times for characters with predictable phonetic radical were significantly shorter than for unpredictable ones, $t = 5.88$, $p < 0.001$. Conversely, under the

untrained condition, no significant differences were observed in reaction time between characters with predictable and unpredictable phonetic radicals, $t = -1.44$, $p = 0.153$. These outcomes suggest that L2 learners can form orthographic expectations based on auditory learning, leading to distinct reaction times when encountering characters with different predictability levels.

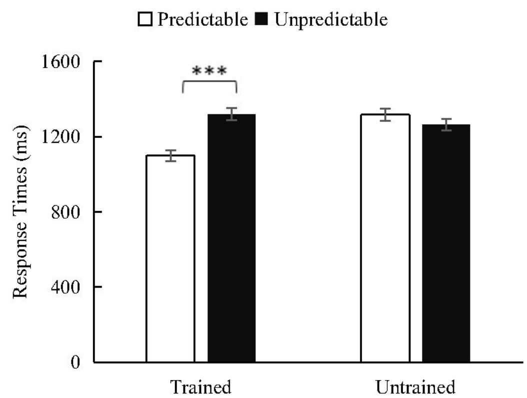

Figure 5 Mean reaction times of phonetic radical predictability under the training conditions.

The Impact of Chinese Writing Proficiency on Orthographic Expectations for Phonogram Characters

Three native Mandarin-speaking graduate students were tasked with scoring the participants' writing of 20 Chinese characters during the writing task. If any rating was inconsistent, it was re-evaluated until all raters reached a consensus. The criteria for scoring were based on whether participants could write characters containing a given phonetic radical that was trained auditorily. Writing the character with the included phonetic radical or directly writing the phonetic radical scored 1 point, otherwise 0 points were assigned. A significant difference existed between the two groups regarding their writing scores, $t(34) = -7.91$, $p < 0.001$. The average reaction times for participants with different writing proficiencies under each condition are presented in Table 5, with statistical results outlined in Table 6.

Table 5 Average reaction times under different conditions for different writing proficiencies

Writing Proficiency	Training Condition	Phonetic Radical Predictability	
		Predictable	Unpredictable
Low Writing Proficiency	Trained	1306 (556)	1487 (577)
	Untrained	1459 (568)	1405 (533)
High Writing Proficiency	Trained	874 (354)	1141 (558)
	Untrained	1163 (565)	1114 (570)

Table 6 Linear mixed model results for different writing proficiencies under various conditions

Effect	Reaction Time			
	b	*SE*	*t*	95%CI
Intercept	7.03	0.05	149.77^{***}	[6.94, 7.12]
P2 vs. P1	0.07	0.03	2.18^{*}	[0.01, 0.13]
S2 vs. S1	0.07	0.02	4.08^{***}	[0.04, 0.10]
W2 vs. W1	−0.31	0.09	-3.40^{**}	[−0.49, −0.13]
P×S	−0.23	0.03	-6.72^{***}	[−0.29, −0.16]
P×W	0.03	0.05	0.52	[−0.07, 0.13]
S×W	0.08	0.03	2.39^{*}	[0.01, 0.15]
P×S×W	−0.10	0.07	−1.42	[−0.23, 0.04]

Note: P represents Phonetic Radical Predictability, S represents Learning, W represents Writing Proficiency. P2 vs. P1 = Unpredictable vs. Predictable Phonetic Radical, S2 vs. S1 = Untrained vs. Trained, W2 vs. W1 = High Writing Proficiency vs. Low Writing Proficiency.

The statistical analysis in lexical recognition task revealed a significant main effect for writing proficiency, $t = -3.40$, $p = 0.002$, participants with high writing proficiency exhibited significantly shorter response times compared to those with low writing proficiency. Since the focus of this experiment was to examine the effects of writing proficiency, we further analyzed the data by splitting participants into either low or high writing proficiency subsets and only conducted statistical analyses regarding writing proficiency. The average reaction times for

participants with low writing proficiency are presented in Table 7, with statistical results outlined in Table 8.

Table 7 Average reaction times under different conditions for the low writing proficiency group

Training Condition	Phonetic Radical Predictability	
	Predictable	Unpredictable
Trained	1388 (566)	1585 (577)
Untrained	1531 (573)	1486 (540)

Table 8 Linear mixed model results for reaction times under different conditions for the low writing proficiency group

	Reaction Time			
Effect	*b*	*SE*	*t*	95%CI
Intercept	7.25	0.05	151.72^{***}	[7.15, 7.34]
Unpredictable vs. Predictable	0.06	0.03	1.34	[−0.03, 0.15]
Untrained vs. Trained	0.02	0.03	0.59	[−0.04, 0.07]
Predictability × Training	−0.18	0.05	-3.24^{**}	[−0.28, −0.07]

Statistical analyses for participants with low writing proficiency (as shown in Figure 6) showed no significant main effects for phonetic radical predictability, $t = 1.34$, $p = 0.195$, or training condition, $t = 0.59$, $p = 0.554$. However, a significant interaction was observed between phonetic radical predictability and training condition, $t = -3.24$, $p = 0.001$. Further analyses revealed that predictability for trained phonetic radical resulted in shorter reaction times compared to unpredictable ones, $t = 3.38$, $p = 0.001$, while no significant difference was found between predictable and unpredictable phonetic radical for untrained condition, $t = -0.61$, $p = 0.542$.

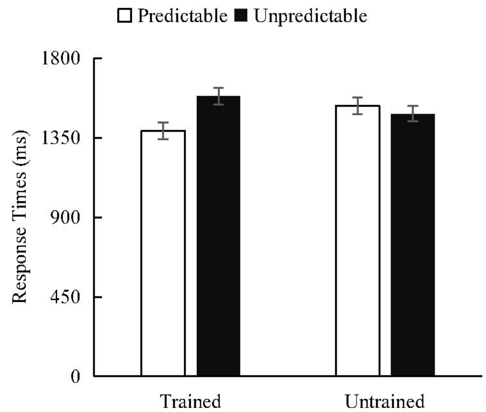

Figure 6 Interaction effects between phonetic radical predictability and training for participants with low writing proficiency.

Table 9 Average reaction times under different conditions for the high writing proficiency group

Training Condition	Phonetic Radical Predictability	
	Predictable	Unpredictable
Trained	874 (333)	1119 (523)
Untrained	1151 (540)	1097 (534)

Table 10 Linear mixed model results for reaction times under different conditions for the high writing proficiency group

Effect	Reaction Time			
	b	*SE*	*t*	95%CI
Intercept	6.87	0.07	102.41^{***}	[6.74, 7.00]
Unpredictable vs. Predictable	0.08	0.04	2.13^{*}	[0.01, 0.15]
Untrained vs. Trained	0.11	0.03	3.83^{**}	[0.05, 0.16]
Predictability ×Training	−0.26	0.04	-6.14^{***}	[−0.35, −0.18]

The average reaction times for participants with high writing proficiency are presented in Table 9, with statistical results outlined in Table 10. Statistical analyses for participants with high writing proficiency (as shown in Figure 7) indicated a significant main effect for phonetic radical predictability, t = 2.13, p = 0.044, illustrating shorter reaction times for predictable phonetic

radical compared to unpredictable ones. Additionally, a significant main effect for training was observed, $t = 3.83$, $p = 0.001$, demonstrating shorter reaction times for trained items compared to untrained ones. Moreover, a significant interaction between phonetic radical predictability and training was found, $t = -6.14$, $p < 0.001$. Follow-up analyses revealed that the reaction time for trained predictable phonetic radical was significantly shorter than for unpredictable ones, $t = 6.15$, $p < 0.001$. However, under the untrained condition, no significant difference was observed between predictable and unpredictable phonetic radical, $t = -1.67$, $p = 0.097$.

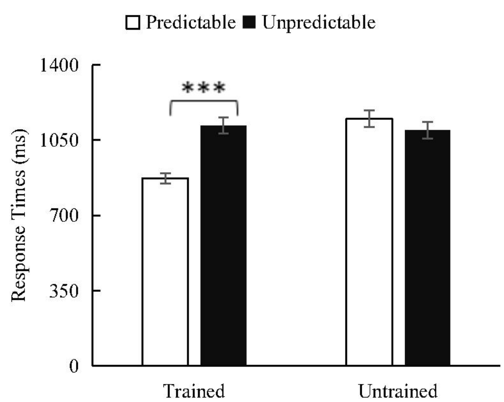

Figure 7 Mean reaction times for participants with high writing proficiency.

Discussion

The aim of Experiment 2 was to further investigate whether the predictability of phonetic radicals and Chinese writing proficiency influence the formation of orthographic expectancy for phonogram characters. The findings regarding phonetic radical predictability indicated that characters trained during the experiment had shorter recognition time compared to those untrained, and characters with predictable phonetic radical had shorter recognition time

compared to those with unpredictable phonetic radical. Crucially, a significant interaction effect was observed between phonetic radical predictability and training. Further analyses suggested a stronger influence of phonetic radical predictability on phonogram characters trained through auditory learning than on those untrained. The results reveal that L2 learners exhibit varied learning outcomes when encountering the written forms of newly trained characters for the first time after learning their oral forms. This pattern hinges on whether the phonetic radical in the characters could be predicted from the trained auditory information: when phonetic radical information was predictable (i.e., aligning with their expectations formed during auditory learning), participants spent less time recognizing the phonogram. Conversely, when phonetic radical information was unpredictable, participants required more time for recognition due to the disparity between their predictions and the actual written forms. However, for phonogram characters untrained through auditory learning, no significant difference was observed in reaction times between predictable and unpredictable phonetic radical conditions due to the lack of expectations.

Learners' performance on the Chinese character writing task is also shown to be an important factor. The results showed that the reaction times of the high writing proficiency group were significantly shorter than those of the low writing proficiency group. This indicates that proficiency in Chinese character writing among L2 learners indeed impacts the generation of orthographic representations. Further analysis of different writing proficiency levels revealed a significant interaction between phonetic radical predictability and learning conditions for both low and high writing proficiency learners. Specifically, reaction times were shorter for *trained* phonogram characters with predictable phonetic radicals compared to unpredictable phonetic radicals. However, under the *untrained* condition, there was no significant difference between

phonogram characters with predictably and unpredictably phonetic radicals. In addition, the high writing proficiency group yielded significant main effects for both phonetic predictability and training condition. Specifically, phonograms trained through auditory practice had shorter reaction times compared to untrained ones, and characters with predictably written phonetic radicals yielded shorter reaction time than unpredictably written phonogram characters. While both groups exhibited interaction effects, the underlying reasons for these effects, when combined with the participants' performance in the writing task, might differ. The selected phonogram characters in the materials were familiar to the L2 learners. However, low writing proficiency individuals did not exhibit errors such as missing or adding strokes during the writing task. Instead, their performance reflected an inability to write Chinese characters or producing characters that did not correspond to actual Chinese characters. This directly indicates their lack of orthographic expectations for trained phonogram characters, suggesting that low writing proficiency L2 learners might not rely on orthographic expectations formed from learning new phonogram characters to make recognition judgments. In contrast, individuals with high writing proficiency demonstrated proficient performance in the writing task, indicating their ability to form orthographic expectations for newly trained phonogram characters. This result pattern is similar to that found in Wegener et al. (2018).

In sum, these results exhibited an interaction between training condition and phonetic radical predictability, akin to the interaction found by Wegner et al. (2020) in English lexical recognition task. This aligns with the *orthographic skeletons hypothesis*, suggesting that Chinese L2 learners, through auditory learning, can employ phonology-orthography correspondences to form orthographic expectations or representations for newly trained characters. However, in

complex writing systems like Chinese (Chang et al., 2016), learners may struggle without sufficient familiarity or an accurate understanding of character representation. They might only be able to map characters from their visual form to their pronunciation. As a result, they fail to develop the reverse mapping from sound to form. Consequently, they may fail to gain accurate orthographic representations for newly trained phonogram characters. However, learners with a certain level of proficiency in Chinese character writing might have accumulated and internalized relevant knowledge about Chinese orthographic rules (such as positional and phonetic compound rules), thus enabling them to form orthographic expectations through sound-to-form mapping.

General Discussion

This study utilized a novel word learning paradigm and lexical recognition task to examine whether L2 learners of Chinese could generate orthographic representations of newly learned Chinese characters solely based on auditory learning through sound-to-form correspondences. Comprehensive analyses of the two experiments reveals that, similar to alphabetic languages, L2 learners can form orthographic expectations for phonogram characters through the phonology-orthography correspondence present in the phonetic radicals. However, due to the differences in the phonology-orthography correspondence units between Chinese characters and alphabetic languages, only L2 learners with a certain level of Chinese writing proficiency can form more accurate orthographic expectations for phonogram characters. Additionally, under the conditions of this experiment, no evidence was found that learners could form orthographic expectations for non-phonogram characters.

The preliminary exploration results of different types of Chinese characters in this study are not the same. The reason for this outcome might be related to the transparency of orthography among different types of Chinese characters. Bakhtiar et al. (2021) conducted a study with 144 children whose native language was Persian. Using words with transparent and opaque spelling-sound relationships, they found that third-grade children recognized transparent word forms faster than opaque ones, suggesting that learning opaque word forms incurs certain processing costs. In Chinese, non-phonogram characters inherently lack phonetic clues and clear regularities. During the learning process, it is therefore difficult for learners to establish a mutually accessible relationship between form and sound. As a result, they may primarily encode new characters at a holistic level through rote memorization (Tseng et al., 2023). Previous research on learning strategies among international students revealed that during the initial stages of learning Chinese, students heavily rely on holistic character-form strategies (Zhao & Jiang, 2002), involving mechanical copying of character forms or memorizing characters as wholes, lacking interconnected exercises between sound, form, and meaning. However, for phonogram characters, the existing correspondence makes them more regular and transparent, enabling learners to form expectations based on this relationship.

Apart from this, the reasons for these findings might also be related to language distance. Dong et al. (2021) utilized representational similarity analysis to estimate neural pattern similarities among Uyghur-Chinese-English trilinguals in their native language and two non-native languages. Behavioral results indicated that participants responded faster in a lexical naming task to English words more similar to their native language. All participants in the current study were speakers of languages written with an alphabetic system, implying that for

them, the similarity between phonogram characters and their native language is greater than that of non-phonogram characters. And previous research has found that L2 word recognition processes are influenced by native language features (Tong et al., 2016). Therefore, during the process of learning phonogram characters, L2 learners may transfer common learning strategies from their native language to the second language, thereby being more adept at utilizing phonetic radicals to provide phonetic information. For non-phonogram characters, being more distant from their native language, tend to be learned using lexical-level representations (Lemhöfer et al., 2008).

It is noteworthy that this study found L2 learners capable of forming orthographic expectations for phonogram character through auditory learning, consistent with findings in alphabetic languages (Wegener et al., 2018; Wegener et al., 2023). However, compared to results from alphabetic languages, the observed interaction pattern between phonetic radical predictability and training condition in this study is contrary to Spanish but similar to English and French. Specifically, in Spanish, longer reading time is observed for trained words with unpredictable spelling (Jevtović et al., 2022), whereas in English and French, shorter reading times are observed for predictably spelled, trained words (Jevtović et al., 2023; Wegener et al., 2018). One possible reason is that Spanish is more transparent, having consistent form-sound correspondences. Consequently, Spanish learners rarely encounter inconsistencies or unexpected spellings, making them more sensitive to situations that do not conform to expectations (Jevtović et al, 2022). In contrast to Spanish, English and French containing both consistent and inconsistent form-sound correspondences, resulting in higher uncertainty for learners while reading and spelling unfamiliar words (Ziegler et al., 1996). Similarly, Chinese is not only a non-alphabetic language, it is also an opaque writing system with a deep orthography (Feng et al.,

2001). This leads to higher uncertainty among L2 learners regarding unfamiliar Chinese characters. Consequently, learners are more sensitive to trained characters with phonetic radical predictability in Chinese, showing a result pattern similar to that observed in French and English.

Combining the results of both types of Chinese characters, it is reasonable to speculate that during the process of learning Chinese, L2 learners may initially accumulate certain Chinese character knowledge through paired associations and gradually realize that some characters can be broken down into components. Following these regularities, they may develop phonetic radical awareness (Luo et al., 2011; Yin & McBride, 2015) and establish syllable-grapheme correspondence for phonogram characters. This aligns with the Dynamic Interactive Model in Chinese reading development (Ye & McBride, 2022), which suggests the transitioning from an initial tendency to learn the arbitrary associations between characters' orthography and sound to the effective use of orthography, phonology, and semantics through analytical strategies in character learning.

The results from high versus low Chinese writing proficiency groups showed that there were differences in the formation of orthographic expectations for phonogram characters among L2 learners with different writing proficiencies, this is also a point where Chinese may differ from alphabetic languages in forming orthographic expectations. For the low writing proficiency group, although there is an interaction between learning conditions and phonetic radical predictability, it can be observed from the writing task that learners, when seeing the characters, can recognize their pronunciation and establish a connection between form and sound. Their inadequacy or lack of mastery in writing Chinese characters might indicate unfamiliarity or inaccuracy in the orthographic representation of Chinese characters, such that they are unable to

mentally form orthographic expectations of Chinese characters during writing task. Consequently, when presented with a character during a recognition task, learners might resort to phonological decoding to obtain its pronunciation (Share, 1995). In this case, if the obtained pronunciation is familiar, they could attempt to match the decoded pronunciation with the one stored in memory, thereby exhibiting facilitation for predictably trained phonetic radicals. Wegener et al. (2018) also suggested that this mechanism, wherein oral language plays a role in reading, relies on learners using phonological decoding from visual exposure through written words to match it with known oral vocabulary. Similar findings were observed in a cross-sectional study where most first and second-grade children displayed inaccurate or insufficient form-sound knowledge, lacking precise knowledge of each letter's direction and position needed for accurate writing. This affected their ability to correctly write letters, subsequently impacting their recognition and reading abilities (Mathwin et al., 2023). For learners with high writing proficiency, they might also recognize new words by storing their pronunciation in memory. However, their performance in writing task clearly demonstrates their ability to form orthographic expectations for the specific phonogram characters, indicating the significance of possessing a certain level of Chinese writing proficiency in forming orthographic expectations for phonogram characters.

Overall, results from the current study suggest that oral language learning can enable intermediate level L2 learners to form orthographic expectations for phonogram characters. However, unlike alphabetic languages, if learners are not sufficiently familiar with the writing rules and orthographic representations of Chinese characters, they may not accurately form orthographic expectations for newly learned phonogram characters through sound-form correspondence. While studies on Chinese and alphabetic languages show these differences, they

all arrive at similar conclusions that learners can generate orthographic expectations related to new written forms through oral language learning. As Jevtović et al. (2022) suggest, the effect appears to be quite robust, having been observed across different groups (children and adults) and techniques (behavioral and eye-tracking). This study has extended this to a logographic language and in L2 learners, indicating that the *orthographic skeleton hypothesis* may have universal applicability. Additionally, the findings of this study align with broader research, supporting the universal influence of orthography on spoken word processing (Ito, 2019; Qu & Damian, 2017; Zou et al., 2012).

In summary, the exploration of orthographic expectations in Chinese in this study contributes to a better understanding of the applicability of the *orthographic skeleton hypothesis* in different languages. Based on these findings, firstly, pedagogical focuses and strategies should be adjusted based on different types of Chinese characters. While non-phonogram characters may not provide pronunciation clues and constitute a smaller proportion, according to McBride (2016), the proportion of these characters is less than 20%, they can help learners form an impression of the overall shape of Chinese characters. Therefore, reinforcing non-phonetic characters' sound-form associations should be emphasized in teaching. Phonogram characters, known for conveying phonetic information, endow their phonetic radicals with predictability. Hence, during the early stages of learning Chinese, introducing the forms and functions of phonetic radical components in phonogram characters could encourage learners to actively retrieve phonetic information from unfamiliar characters, making it easier for foreign learners to grasp Chinese. Secondly, while alphabetic languages require memorizing some letters and their pronunciations in order to read through print and spell when hearing a word, the Chinese writing system is a complex visual language consisting of thousands of logographic units (Yang et al.,

2019). These characters are difficult to write, and should therefore necessitate the pedagogical approach that emphasizes characters' internal structures and orthographic rules. At the same time, increased writing practices and instructions regarding the positional regularities of character components will also help learners acquire characters. This in particular will aid low writing proficiency learners in establishing bidirectional correspondences between sound and form, thereby enhancing their reading skills.

Conclusion and Future Research

Results from this study indicate that in the absence of contextual cues, such as lexical recognition task, Chinese L2 learners can generate orthographic expectations for newly learned phonogram characters, but the accuracy of these expectations is related to their Chinese writing proficiency. This demonstrates the presence of the orthographic expectation effect in Chinese and provides further support for the *orthographic skeleton hypothesis*. However, under the current experimental conditions, no evidence was found that L2 learners can generate orthographic expectations for newly learned non-phonogram characters.

While this study offers several valuable insights, it is essential to acknowledge its limitations when interpreting the findings. First, the present study did not make a specific distinction between L1 types of L2 learners, therefore future research should further distinguish whether L2 learners with varying language distances from Chinese exhibit differences in the formation of orthographic expectations to validate this point. Second, semantic radicals were used in this study that cannot independently form characters. In fact, 58% of the Chinese characters learned provide conceptual categories or are directly related to the meaning of the

characters (Shu et al., 2003), namely semantic radical, which can also have the characteristics of conveying both meaning and form. Hence, future studies could also further explore whether learners can form orthographic expectations through semantic-form associations for semantic radicals.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bakhtiar, M., Mokhlesin, M., Pattamadilok, C., Politzer-Ahles, S., & Zhang, C. (2021). The effect of orthographic transparency on auditory word recognition across the development of reading proficiency. *Frontiers in Psychology, 12*, 691989. https://doi.org/10.3389/fpsyg.2021.691989

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., M., Maechler, M., Bolker, B., & Walker, S. (2018). *lme4: Linear mixed-effects models using 'Eigen' and S4*. Version 1.1-31. https://cran.r-project.org/web/packages/lme4/index.html

Beyersmann, E., Wegener, S., Spencer, J., & Castles, A. (2023). Acquisition of orthographic forms via spoken complex word training. *Psychonomic Bulletin & Review, 30*(2), 739–750. https://doi.org/10.3758/s13423-022-02185-y

Chang, L. Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading, 20*(1), 64–85. https://doi.org/10.1080/10888438.2015.1104688

Chen, L. I. N., Perfetti, C. A., & Leng, Y. (2019). Reading Pinyin activates character orthography for highly experienced learners of Chinese. *Bilingualism: Language and Cognition*, *22*(1), 103-111. https://doi.org/10.1017/S136672891700058X

Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon, 8*(3), 372–420. https://doi.org/10.1075/ml.8.3.06dar

Dong, J., Li, A., Chen, C., Qu, J., Jiang, N., Sun, Y., Hu, L., & Mei, L. (2021). Language distance in orthographic transparency affects cross - language pattern similarity between native and non - native languages. *Human Brain Mapping, 42*(4), 893–907. https://doi.org/10.1002/hbm.25266

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading, 16*(6), 504–525. https://doi.org/10.1080/10888438.2011.598199

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feng, G., Miller, K., Shu, H., & Zhang, H. (2001). Rowed to recovery: The use of phonological and orthographic information in reading Chinese and English. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(4), 1079–1100. https://doi.org/10.1037/0278-7393.27.4.1079

He, Y., Wang, Q., & Anderson, R. C. (2005). Chinese children's use of subcharacter information about pronunciation. *Journal of Educational Psychology, 97*(4), 572–579. https://doi.org/10.1037/0022-0663.97.4.572

Hsiao, J. H., & Liu, T. (2010). Position of phonetic components may influence how written words are processed in the brain: Evidence from Chinese phonetic compound pronunciation. *Cognitive, Affective, & Behavioral Neuroscience, 10*(4), 552–559. https://doi.org/10.3758/cabn.10.4.552

Hsieh, C. Y., Lin, W. C., Li, M. F., & Wu, J. T. (2021). The concept of neighborhood underlying the phonetic consistency effect in Chinese character recognition: A four-decade review and beyond. *Reading and Writing, 34*(5), 1225–1246. https://doi.org/10.1007/s11145-020-10119-0

Ito, A. (2019). Prediction of orthographic information during listening comprehension: A printed-word visual world study. *Quarterly Journal of Experimental Psychology, 72*(11), 2584–2596. https://doi.org/10.1177/1747021819851394

Jevtović, M., Antzaka, A., & Martin, C. D. (2022). Gepo with a G, or Jepo with a J? Skilled readers generate orthographic expectations for novel spoken words even when spelling is uncertain. *Cognitive Science, 46*(3), e13118. https://doi.org/10.1111/cogs.13118

Jevtović, M., Antzaka, A., & Martin, C. D. (2023). Déjà-lu: When orthographic representations are generated in the absence of orthography. *Journal of Cognition, 6*(1), 1–20. https://doi.org/10.5334/joc.250

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47–77. https://doi.org/10.1093/applin/21.1.47

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: a megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 12–31. https://doi.org/10.1037/0278-7393.34.1.12

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Liang, Y. (2019). Acquisition of Chinese characters: An error analysis of character writing among nonnative learners. *Language Teaching and Linguistic Studies*, (4), 33–44. https://doi.org/10.3969/j.issn.0257-9448.2019.04.004

Luo, Y. C., Chen, X., Hélène Deacon, S., & Li, H. (2011). Development of Chinese orthographic processing: A cross-cultural perspective. *Writing Systems Research, 3*(1), 69-86. https://doi.org/10.1093/wsr/wsr008

Marian, V., Bartolotti, J., Daniel, N. L., & Hayakawa, S. (2021). Spoken words activate native and non-native letter-to-sound mappings: Evidence from eye tracking. *Brain and Language, 223*, 105045. https://doi.org/10.1016/j.bandl.2021.105045

Mathwin, K., Chapparo, C., Challita, J., & Hinitt, J. (2023). Cross-sectional examination of the proficiency of year 1 and year 2 children's alphabet-letter-writing skills. *Reading and Writing*, 1–24. https://doi.org/10.1007/s11145-023-10476-6

McBride, C. A. (2016). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review, 28*(3), 523–549. https://doi.org/10.1007/s10648-015-9318-2

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition, 80*(3), 231–262. https://doi.org/10.1016/S0010-0277(00)00150-5

Nation, K., & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology, 103*(3), 296–308. https://doi.org/10.1016/j.jecp.2009.03.004

Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J., Bolger, D. J., & Tan, L. H. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition*, 10(2), 131–146. https://doi.org/10.1017/S1366728907002891

Qu, Q., & Damian, M. F. (2017). Orthographic effects in spoken word recognition: Evidence from Chinese. *Psychonomic Bulletin & Review, 24*(3), 901–906. https://doi.org/10.3758/s13423-016-1164-9

Qu, Q., Cui, Z., & Damian, M. F. (2018). Orthographic effects in second-language spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(8), 1325–1332. https://doi.org/10.1037/xlm0000520

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74*(1), 27–47. https://doi.org/10.1111/1467-8624.00519

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30*(2), 139–181. https://doi.org/10.1016/0010-0277(88)90038-8

Tong, X., Kwan, J. L. Y., Wong, D. W. M., Lee, S. M. K., & Yip, J. H. Y. (2016). Toward a dynamic interactive model of non-native Chinese character processing. *Journal of Educational Psychology, 108*(5), 680–693. https://doi.org/10.1037/edu0000083

Tseng, C. C., Hu, J. F., Chang, L. Y., & Chen, H. C. (2023). Learning to read Chinese: The roles of phonological awareness, paired–associate learning, and phonetic radical awareness. *Reading and Writing, 36*(7), 1769–1795. https://doi.org/10.1007/s11145-022-10352-9

Veivo, O., Porretta, V., Hyönä, J., & Järvikivi, J. (2018). Spoken second language words activate native language orthographic information in late second language learners. *Applied Psycholinguistics, 39*(5), 1011–1032. https://doi.org/10.1017/S0142716418000103

Wegener, S., Wang, H. C., Beyersmann, E., Nation, K., Colenbrander, D., & Castles, A. (2023). Orthographic expectancies in the absence of contextual support. *Scientific Studies of Reading, 27*(2), 187–197. https://doi.org/10.1080/10888438.2022.2127356

Wegener, S., Wang, H. C., de Lissa, P., Robidoux, S., Nation, K., & Castles, A. (2018). Children reading spoken words: Interactions between vocabulary and orthographic expectancy. *Developmental Science, 21*(3), e12577. https://doi.org/10.1111/desc.12577

Yan, G., Liu, N., Liang, F., Liu, Z., & Bai, X. (2015). The comparison of eye movements between Chinese children and adults when reading disappearing text. *Acta Psychologica Sinica*. *47*(3), 300–318. https://doi.org/10.3724/SP.J.1041.2015.00300

Yang, Q., Wang, Y., & Zhang, J. J. (2019). Effects of orthographic depth on Chinese word naming for Han and Uyghur students. *Acta Psychologica Sinica, 51*(1), 1–13. https://doi.org/10.3724/SP.J.1041.2019.00001

Ye, Y., & McBride, C. (2022). A dynamic interactive model of Chinese spelling development. *Educational Psychology Review*, *34*, 2897–2917. https://doi.org/10.1007/s10648-022-09684-3

Yin, L., & McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychological Science, 26*(4), 424–432. https://doi.org/10.1177/0956797614567203

Yum, Y. N., Law, S. P., Su, I. F., Lau, K. Y. D., & Mo, K. N. (2014). An ERP study of effects of regularity and consistency in delayed naming and lexicality judgment in a logographic writing system. *Frontiers in Psychology, 5*, 315–327. https://doi.org/10.3389/fpsyg.2014.00315

Zhang, H., & Roberts, L. (2019). The role of phonological awareness and phonetic radical awareness in acquiring Chinese literacy skills in learners of Chinese as a second language. *System, 81*, 163–178. https://doi.org/10.1016/j.system.2019.02.007

Zhang, J. J., Wang, J., & Yin, C. (2014). The role of phonetic radicals and semantic radicals in phonetics and semantics extraction of phonogram characters: An eye movement study on components perception. *Acta Psychologica Sinica, 46*(7), 885–900. https://doi.org/10.3724/SP.J.1041.2014.00885

Zhang, J., Li, H., & Liu, Y. (2021). The influence of orthography on oral vocabulary acquisition in learners of Chinese as a second language. *Studies in Second Language Acquisition, 43*(5), 1157–1172. https://doi.org/10.1017/S0272263121000139

Zhao, G., & Jiang X. (2002). What is the most effective strategy for learning Chinese characters: A survey among CSL beginners. *Applied Linguistics*, (2), 79–85. https://doi.org/10.16499/j.cnki.1003-5397.2002.02.014

Zhou, L., Duff, F. J., & Hulme, C. (2015). Phonological and semantic knowledge are causal influences on learning to read words in Chinese. *Scientific Studies of Reading, 19*(6), 409–418. https://doi.org/10.1080/10888438.2015.1068317

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers, 28*(4), 504–515. https://doi.org/10.3758/BF03200539

Zou, L., Desroches, A. S., Liu, Y., Xia, Z., & Shu, H. (2012). Orthographic facilitation in Chinese spoken word recognition: An ERP study. *Brain and Language, 123*(3), 164–173. https://doi.org/10.1016/j.bandl.2012.09.006